

**REMARKS**

In the Office Action, the Examiner rejected claims 1 and 7 under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,463,618 (hereinafter, "Furukawa") in view of so-called Applicant's Admitted Prior Art (hereinafter, "AAPA"); rejected claims 2 and 8 under 35 U.S.C. § 103(a) as being unpatentable over Furukawa in view of AAPA and further in view of U.S. Patent No. 5,475,791 (hereinafter "Schalk"); rejected claims 3 and 9 under 35 U.S.C. § 103(a) as being unpatentable over Furukawa, AAPA, Schalk and further in view of "Continuous Speech Recognition in Noise Using Spectral Subtraction and HMM Adaptation," 1994 (hereinafter, "Flores"); rejected claims 4, 5, 10, and 11 under 35 U.S.C. § 103(a) as being unpatentable over Furukawa, AAPA, Schalk and further in view of "Signal Conditioning Techniques for Robust Speech Recognition," 1996 (hereinafter, "Rahim"), Flores, and so-called well known prior art; and rejected claims 6 and 12 under 35 U.S.C. § 103(a) as being unpatentable over Rahim and so-called well known prior art.

By this Amendment, Applicants amend claims 1 and 7. Claims 1-12 are pending in the present application. Applicants respectfully traverse the above-noted rejections of claims 1-12 for the following reasons.

**The Rejection to Claims 1 and 7 under 35 U.S.C. § 103(a) Should be Withdrawn**

Applicants respectfully traverse the rejection to claims 1 and 7 under 35 U.S.C. § 103(a) as being rendered obvious by Furukawa and AAPA. Amended claim 1 recites "decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled, said decision

means outputting a result whether said voice is included in the near-end speech signal”

Claim 1 (emphasis added).

The Examiner admits that Furukawa “does not specifically disclose using time and frequency domain information of the acoustic echo-canceled signal for checking whether or not voice is included in the microphone input signal, wherein the microphone input signal comprises background noise.” The Examiner contends, however, that “checking, in each frame, whether or not voice is included in an input signal, by using time domain and frequency domain information of an acoustic echo-canceled signal . . . is known in the art, as indicated by applicant’s admitted prior art at page 20, lines 16-24 of the specification.” Office Action, pg. 3. Applicants disagree.

The specification describes Japanese Patent Application Laid Open No. 213946/1997 as disclosing “an echo canceler which uses a double talk detection circuit that decides whether a speech is included in an input voice signal by using time and frequency domain information of source information on the input voice signal (a signal before being echo-canceled) and on an additive noise with a known source.” Specification, pg. 20, ll. 16-24 (emphasis added). Thus, the input voice signal (the near end speech signal) is not echo-canceled, and therefore, the cited portion of the specification does not teach or suggest each and every element of claim 1 including “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled, said decision means outputting a result whether said voice is included in the near-end speech signal.” Since the Examiner apparently

acknowledges that Furukawa also fails to teach the claimed decision means, Applicants submit that amended claim 1 is allowable over the applied references.

Claim 7 recites a speech processing method including “a decision step for checking . . . .” Claim 7 is therefore similar to claim 1 in this respect and is allowable over Furukawa and AAPA for at least the reasons discussed above with regard to claim 1.

Further, Furukawa concerns “far-end speech” (Furukawa, col. 6, ll.6-13), and thus necessarily fails to teach “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal . . . .” (emphasis added), as recited in claim 1, as well as the decision step recited in claim 7. Therefore, the rejection of claims 1 and 7 should be withdrawn for this reason as well.

In addition, amended claim 1 recites “update means . . . for supplying the updated impulse response to said supply means.” Claim 7 similarly recites an update step . . . for supplying the updated impulse response to said supply step. The Examiner contends that an “updating step” is disclosed by Furukawa at col. 5, ll. 54-56. Office Action, pg. 3. However, the cited portion of Furukawa merely discloses that the “signal processing in the following description is performed all in digital form and that the signal is sampled at a frequency of 8kHz.” Furukawa, col. 5, ll. 54-56. Such vague disclosure of “signal processing” fails to teach or suggest Applicants' claimed step of “updating . . . for continually updating the impulse response for each sample by using said source signal, said acoustic echo-canceled signal and the current impulse response held by the supply step and for supplying the updated impulse response to said supply step,” as recited in claim 7, or “update means,” as recited in claim 1. Thus, claims 1 and 7 are allowable over Furukawa and AAPA for this reason also.

**The Rejection to Claims 2 and 8 under 35 U.S.C. § 103(a) Should be Withdrawn**

Applicants respectfully traverse the rejection to claims 2 and 8 under 35 U.S.C. § 103(a) as being obvious over Furukawa and AAPA and further in view of Schalk. Claim 2 depends upon claim 1 and therefore contains all the limitations contained therein including the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled.” Claim 8 includes similar claim language in light of its dependency from claim 7.

As noted above, Furukawa and AAPA, taken either alone or in combination, do not disclose each and every element of claim 1 including the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information after said acoustic signal has been echo-canceled.” Schalk fails to overcome these deficiencies.

Schalk discloses a “method for recognizing a spoken word in the presence of interfering speech, such as a system-generated voice prompt.” Schalk, Abstract. Schalk provides “‘on-the-fly’ switching between echo cancellation and recognition processes.” Schalk, col. 3, ll. 41-43. To detect a voice, Schalk processes energy from a signal to determine an onset of the spoken word by calculating a root mean square value of the signal and comparing the value to a threshold value. Schalk, col. 4, ll. 32-36. Schalk conducts speech recognition after detecting a voice or a “word onset.” Schalk, col. 4, ll. 28-32. Upon detection of a word onset, speech recognition may occur by activating a recognizer. Schalk, col. 4, ll. 45-53. Schalk, however, switches between echo-canceling and speech recognition. Schalk, col. 3, ll. 41-45. Therefore, echo-cancelling does not

occur during speech recognition or further detection of speech, because Schalk “obviates dedicating a portion of the digital signal processor to provide echo cancellation during the entire spoken utterance.” Schalk, col. 5, ll. 59-61. Therefore, a portion of the signal in Schalk is not echo-canceled but is processed for detection of speech and recognition of that speech. Accordingly, Schalk does not teach or suggest each and every element of claim 1, including “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled (emphasis added).” Schalk, therefore, does not overcome the shortcomings of Furukawa and AAPA and claim 2 is allowable at least due to its dependence from claim 1.

Further, Applicants respectfully request the Examiner to withdraw the rejection of claim 8, which contains similar claim language as claim 2, but depends from claim 7 and is directed to a “speech processing method.”

In addition, Schalk does not make any mention of the claimed “update means” of claim 1 and the “update step” of claim 7. Accordingly, the Examiner fails to provide a prima facie case for obviousness of claims 2 and 8 for this reason as well.

**The Rejection to Claims 3 and 9 under 35 U.S.C. § 103(a) Should be Withdrawn**

As noted above, Furukawa, AAPA, and Schalk either alone or in combination, do not disclose each and every element of claim 1, from which claim 3 indirectly depends, including the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency

domain information after said acoustic signal has been echo-canceled.” Flores fails to overcome these deficiencies.

Flores discloses “a scheme for robust speech recognition at very poor signal to noise ratios . . . .” Flores, pg. I-409, Abstract. However, Flores does not disclose, nor does the Examiner argue that Flores teaches, the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information after said acoustic signal has been echo-canceled.”

Therefore, Furukawa, AAPA, Schalk, and Flores, either individually or in combination do not disclose each and every element of claim 3 and, thus, the Examiner has not provided a *prima facie* case for obviousness. Applicants respectfully request the Examiner to withdraw the rejection to claim 3. Further, Applicants respectfully request the Examiner to withdraw the rejection of claim 9, which contains similar claim language as claim 3, but is directed to a “speech processing method.”

In addition, none of the applied references disclose “means for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source,” because each of the references concerns removing noise from known sources, such as noise inputted acoustically from a speaker or microphone. See claim 3 (emphasis added). Claim 9 recites “a step for successively determining a spectrum mean for each frame based on the spectrum obtained; and a step for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source.” Thus, in addition to the reasons set forth, the

applied references, alone or in combination, fail to render obvious claims 3 and 9 for this reason as well.

**The Rejection to Claims 4-5 and 10-11 under 35 U.S.C. § 103(a) Should be Withdrawn**

As noted above, Furukawa, AAPA, Schalk, and Flores either alone or in combination, do not disclose each and every element of claim 1, from which claims 4 and 5 indirectly depend, including the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information after said acoustic signal has been echo-canceled..” Rahim and so-called well known prior art, as allegedly evidenced by O'Shaughnessy, pp. 229-231 or Parsons, pp. 203-205, fail to overcome their deficiencies.

Rahim teaches that an “[a]coustic mismatch encountered in various training and testing conditions of hidden Markov model (HMM) based systems [which] often causes severe degradation in speech recognition performance.” Rahim, pg. 107, Abstract. Rahim discloses three techniques for blind channel equalization, namely, cepstral mean subtraction (CMS), signal bias removal (SBR), and hierarchical signal bias removal (HSBR). *Id.* However, Rahim does not disclose the claimed “decision means for checking, in each frame, whether or not a voice is included in the near-end speech signal, by using time domain information and frequency domain information of said acoustic signal after said acoustic signal has been echo-canceled.” Further, O'Shaughnessy and Parsons also fail to disclose the above-noted claim language, nor does the Examiner rely on Rahim, O'Shaughnessy, and Parsons for such teachings.

Therefore, Furukawa, AAPA, Schalk, Flores, Rahim, O'Shaughnessy, and Parsons taken either individually or in combination do not disclose each and every element of claim 3 and, thus, the Examiner has not provided a *prima facie* case of obviousness. Applicants respectfully request the Examiner to withdraw the rejection to claim 3. In addition, Applicants respectfully request the Examiner to withdraw the rejection of claim 9, which contains similar claim language as claim 3, but is directed to a "speech processing method."

**The Rejection to Claims 6 and 12 under 35 U.S.C. § 103(a) Should be Withdrawn**

Applicants respectfully traverse the rejection to claims 6 and 12 under 35 U.S.C. § 103(a) as being rendered obvious by Rahim and "well-known prior art" for which the Examiner cited U.S. Patent Nos. 6,539,352 (Sharma), 5,742,694 (Eatwell), 6,001,131 (Raman), and 6,263,307 (Arslan) in the Advisory Action of August 6, 2004. Claim 6 recites "means . . . for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone." See claim 6. The Examiner concedes that Rahim does not teach implementing cepstral mean subtraction on a non-speech cepstrum and concludes that it is well-known in the art to provide for estimates of non speech (or noises) in the implementation of a subtraction scheme for noise suppression citing to Sharma, Eatwell, Raman, and Arslan. See Office Action, pg. 7 and Advisory Action dated August 6, 2004, pg. 3. However, the Examiner fails to identify specific teachings in Sharma, Eatwell, Raman and Arslan corresponding to the claimed "means for subtracting" as recited in claim



6. In fact, the Examiner concedes in the last Office Action at pg. 8 that the cited references of “well known prior art” show only estimates of non speech, but do not disclose “means . . . for subtracting the *cepstrum mean of the non-speech frame* of each talker from the cepstrum of the non-speech frame of the talker to correct *multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.*” See claim 6 (emphasis added). Thus, the Examiner has not provided a *prima facie* case for obviousness.

In light of the above-described deficiencies of Rahim, Sharma, Eatwell, Raman, and Arslan, Applicants submit that claim 6 is allowable over the applied references.

Moreover, claim 12 recites limitations similar to that of claim 6. For example, claim 12 recites “a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.” Claim 12, therefore, is allowable for at least the reasons discussed above in connection with claim 6. Accordingly, Applicants respectfully request the Examiner to withdraw the rejection of claim 12.

In view of the foregoing amendments and remarks, Applicants respectfully request reconsideration and reexamination of this application and the timely allowance of the pending claims.

Please grant any extensions of time required to enter this response and charge any additional required fees to our deposit account 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW,  
GARRETT & DUNNER, L.L.P.

*David Sody*

By: \_\_\_\_\_

*Reg. No. 34,731*

*for* John M. Romary  
Reg. No. 26,331

Dated: December 23, 2004